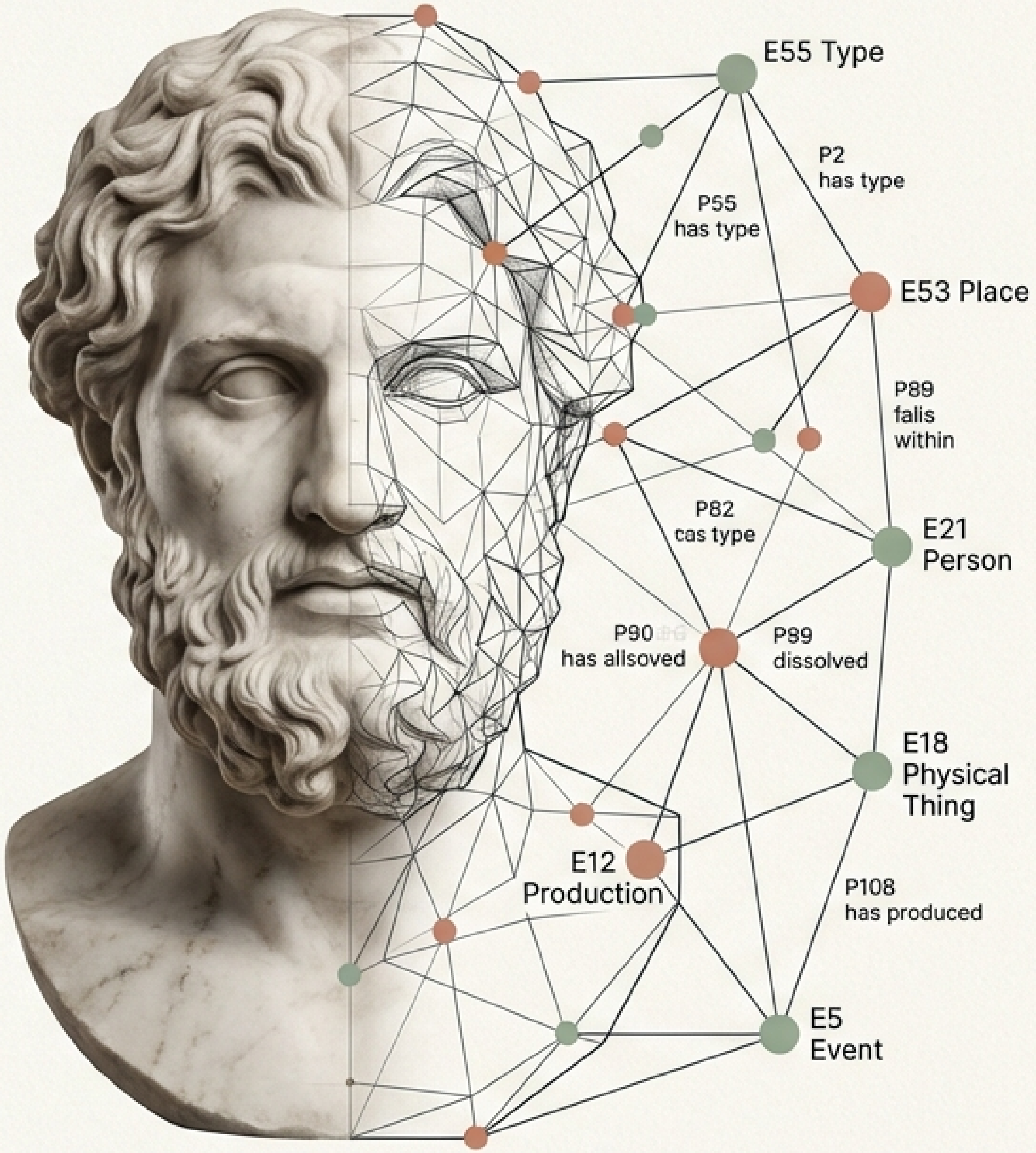
CIDOC CRM: The Semantic Glue for Cultural Heritage

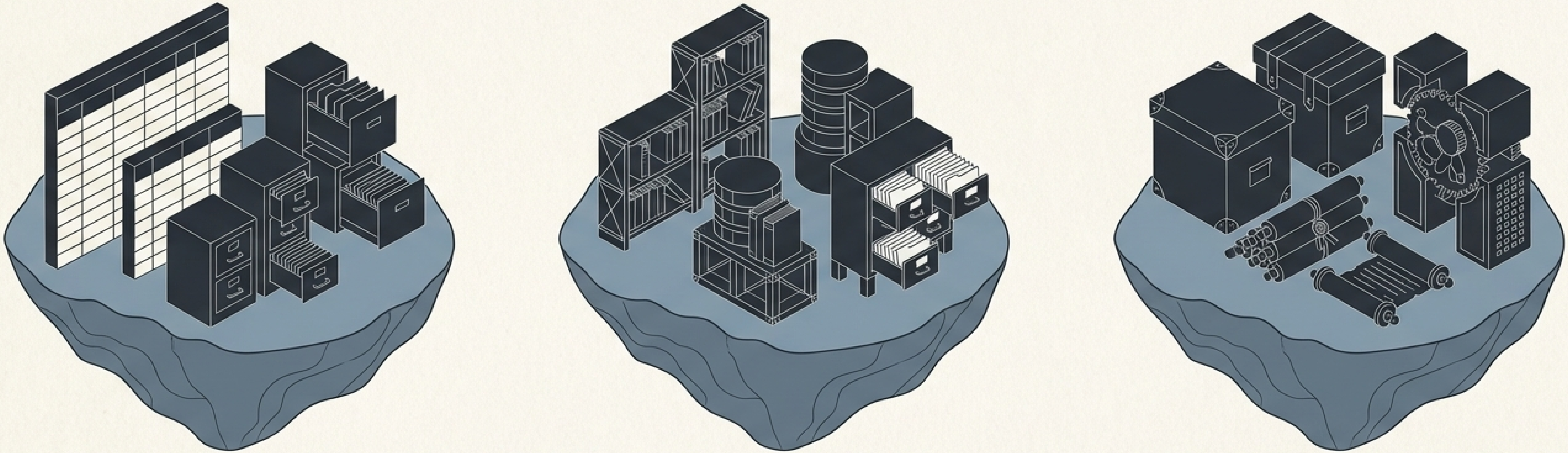
Mapping the human past at a human scale.


E55 Type
P2 has type
P55 has type
E53 Place
P89 falis within
P82 cas type
E21 Person
P90 has allsoved
P99 dissolved
E18 Physical Thing
E12 Production
P108 has produced
E5 Event

The Fragmentation of Historical Data

Cultural heritage data is rich, diverse, and heavily dispersed. Without a shared language, institutions document history using localized schemas and structures. This fragmentation prevents researchers from seamlessly querying, exploring, and integrating knowledge across different domains.

Museums **Libraries** **Archives**

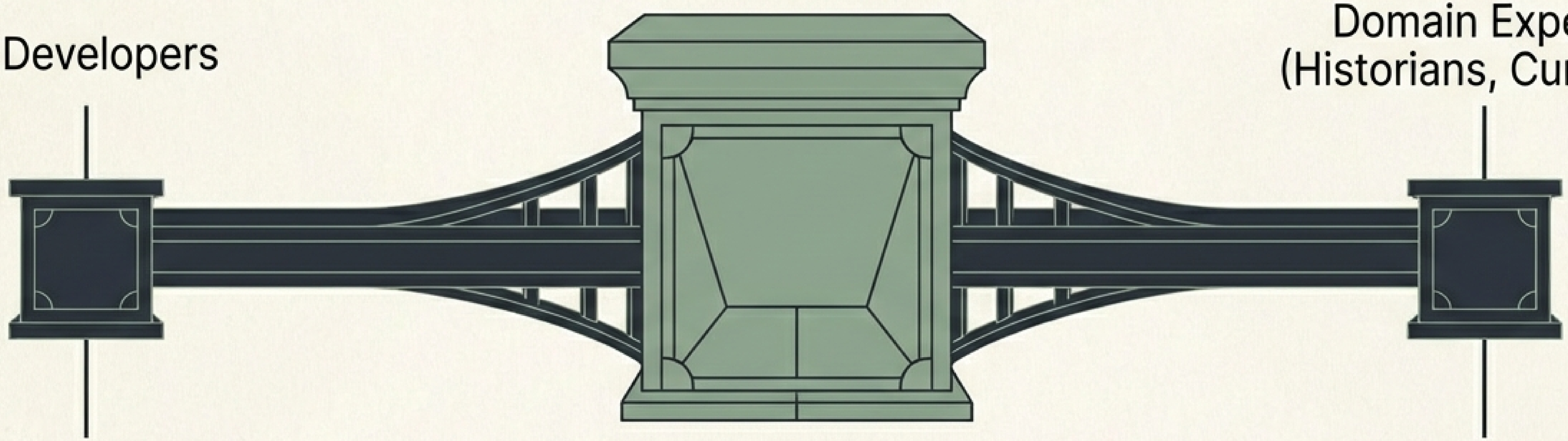


Museums **Libraries** **Archives**

A Formal Ontology for Information Integration

The CIDOC Conceptual Reference Model (CRM) is an official ISO standard and a formal ontology born from over 20 years of collaborative development.

It provides the semantic glue needed to integrate data from multiple sources in a software- and schema-agnostic fashion, **serving as a common language** for both developers and historians.



IT Developers
Domain Experts
(Historians, Curators)

IT Developers

CIDOC CRM
(ISO 21127:2023)

Domain Experts
(Historians, Curators)

Mapping the Past in an Open World

The intended scope of the CRM is **curated, factual knowledge** based on material evidence.

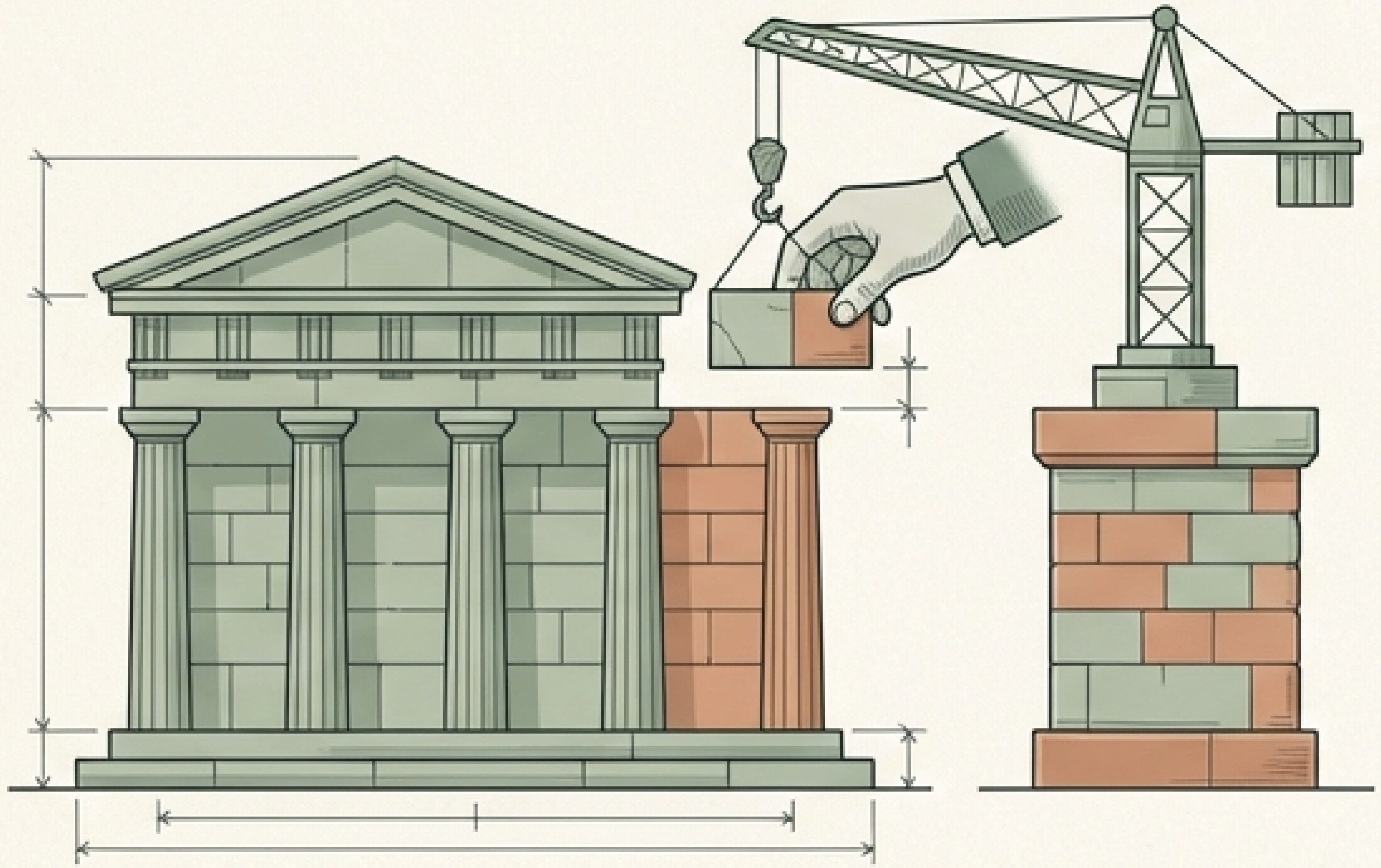
However, it operates strictly on the **Open World Assumption**:

- Our records of the past are inherently incomplete.
- The absence of a property in a database does not mean it does not exist in reality.
- Classes cannot be defined merely by enumerating known instances.

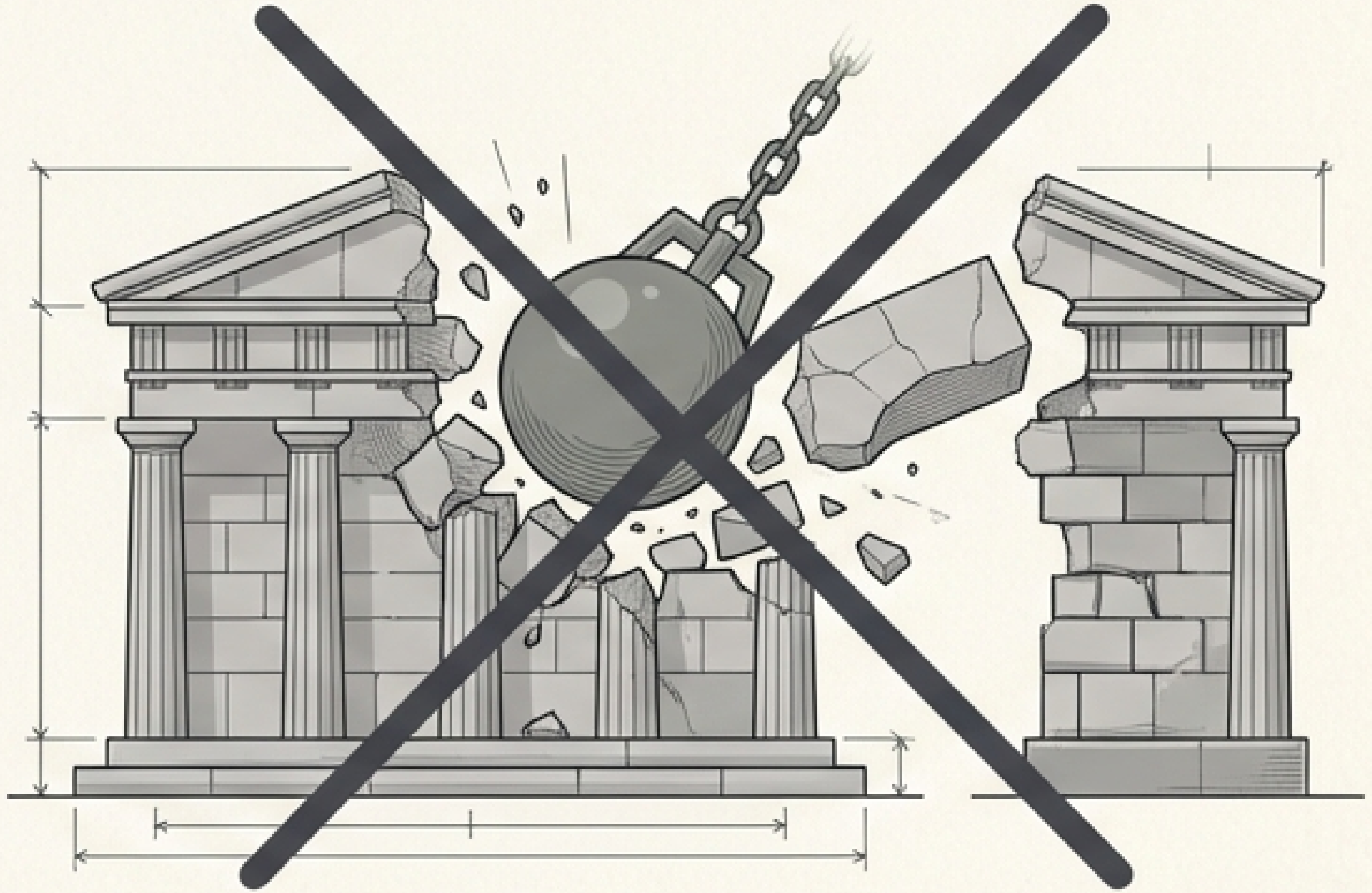
Knowledge Adds Up; It Doesn’t Break

The CIDOC CRM applies the Principle of Monotonicity.

- System never invalidates existing, well-formed deductions.
- New facts refine the model rather than forcing a rewrite.
- The CRM strictly avoids complement classes—knowledge simply accumulates conflict-free.



Monotonicity: Safe Accumulation



Non-Monotonic: Destructive Rewrites

The Anaton Anatomy of Historical Discourse

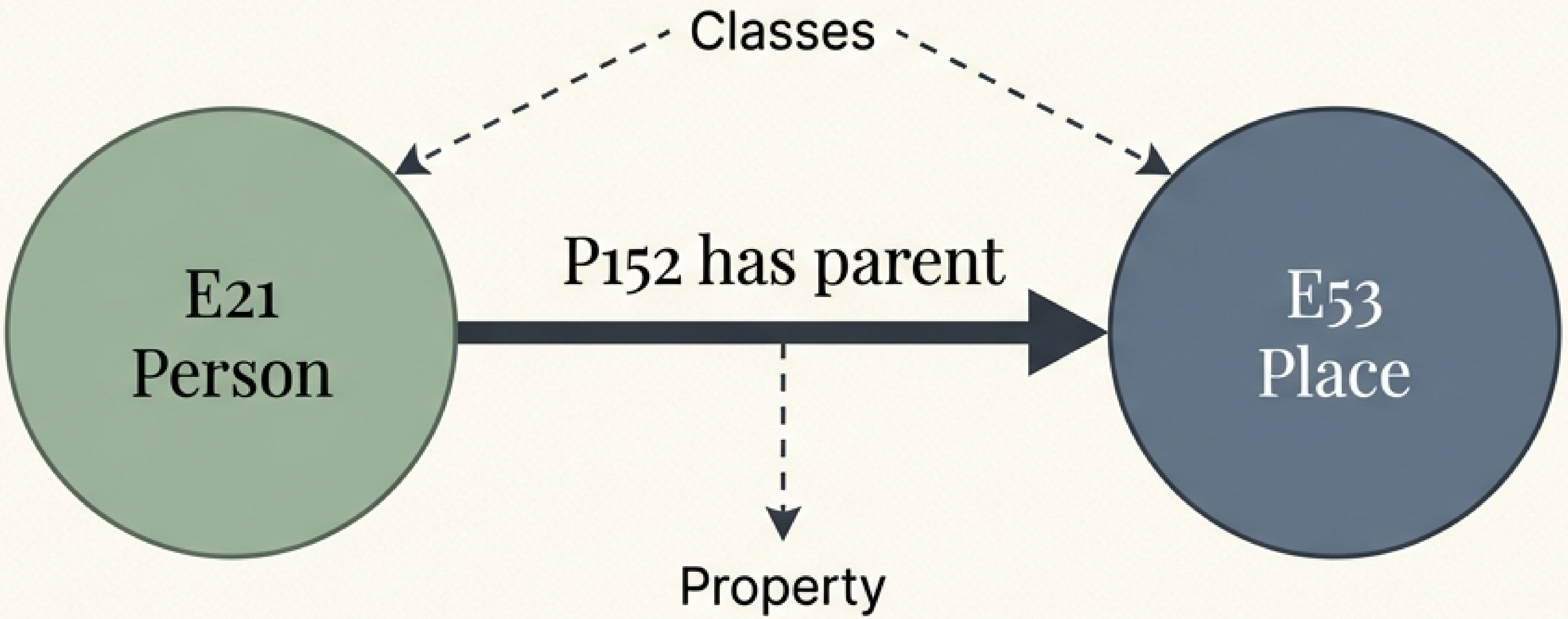
The ontology is built on two primary semantic constructs:

Classes (E-numbers)

The Nouns of the model. They represent general notions in the domain of discourse.

Properties (P-numbers)

The Verbs of the model. They define specific directional relationships between domains and ranges.


Classes
E21
Person
P152 has parent
E53
Place
Property

Two Distinct Ways to Exist in Time

E77 Persistent Items (Endurants)

Things, people, and ideas that endure. Wholly present at any time they exist and preserve their identity between events.

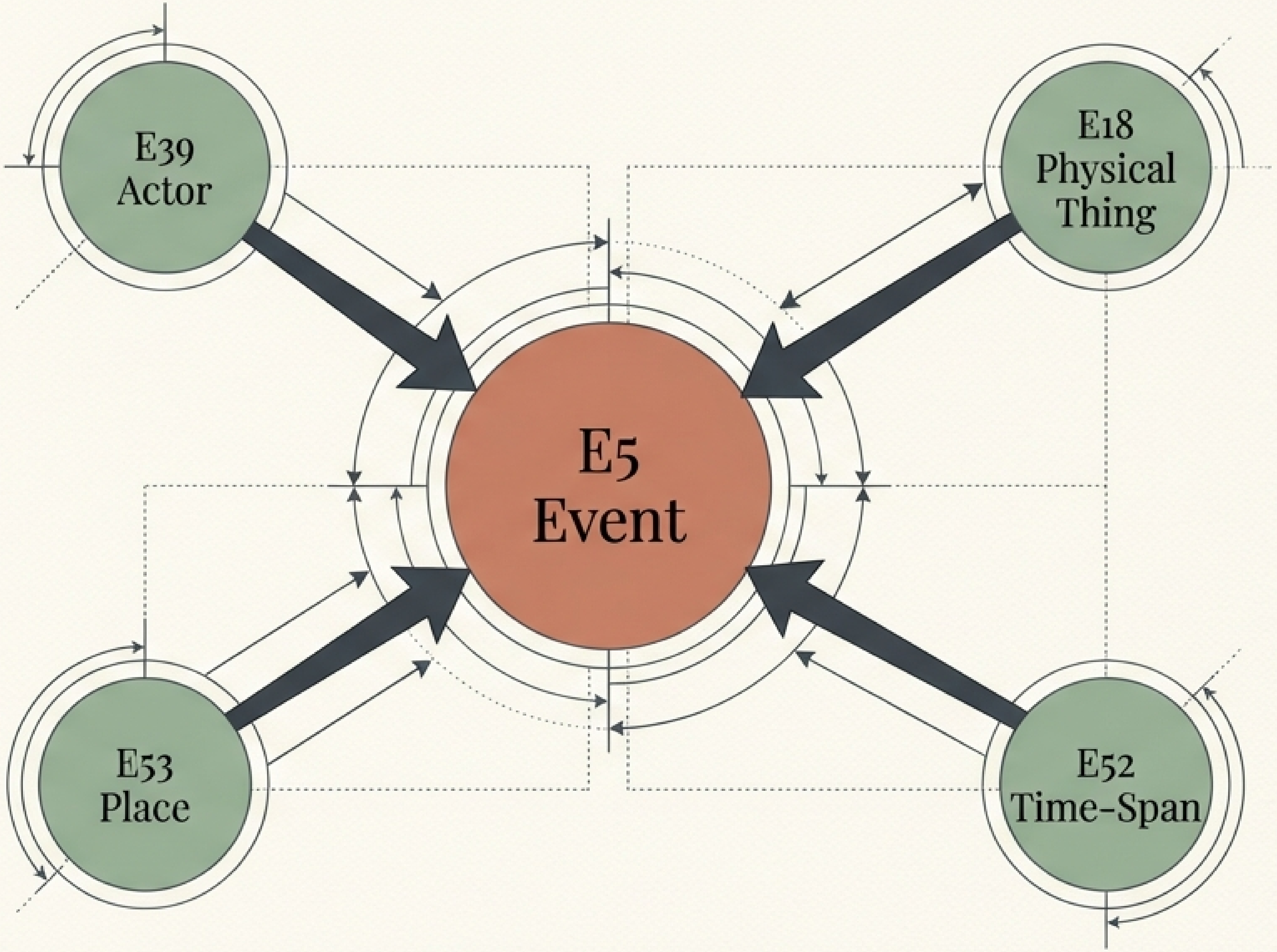
E2 Temporal Entities (Perdurants)

Events and periods that happen in time. They extend in time by accumulating temporal phases and eventually pass.

History is a Network of Meetings

The CIDOC CRM is profoundly event-centric. Persistent items don't independently generate history; they interact.

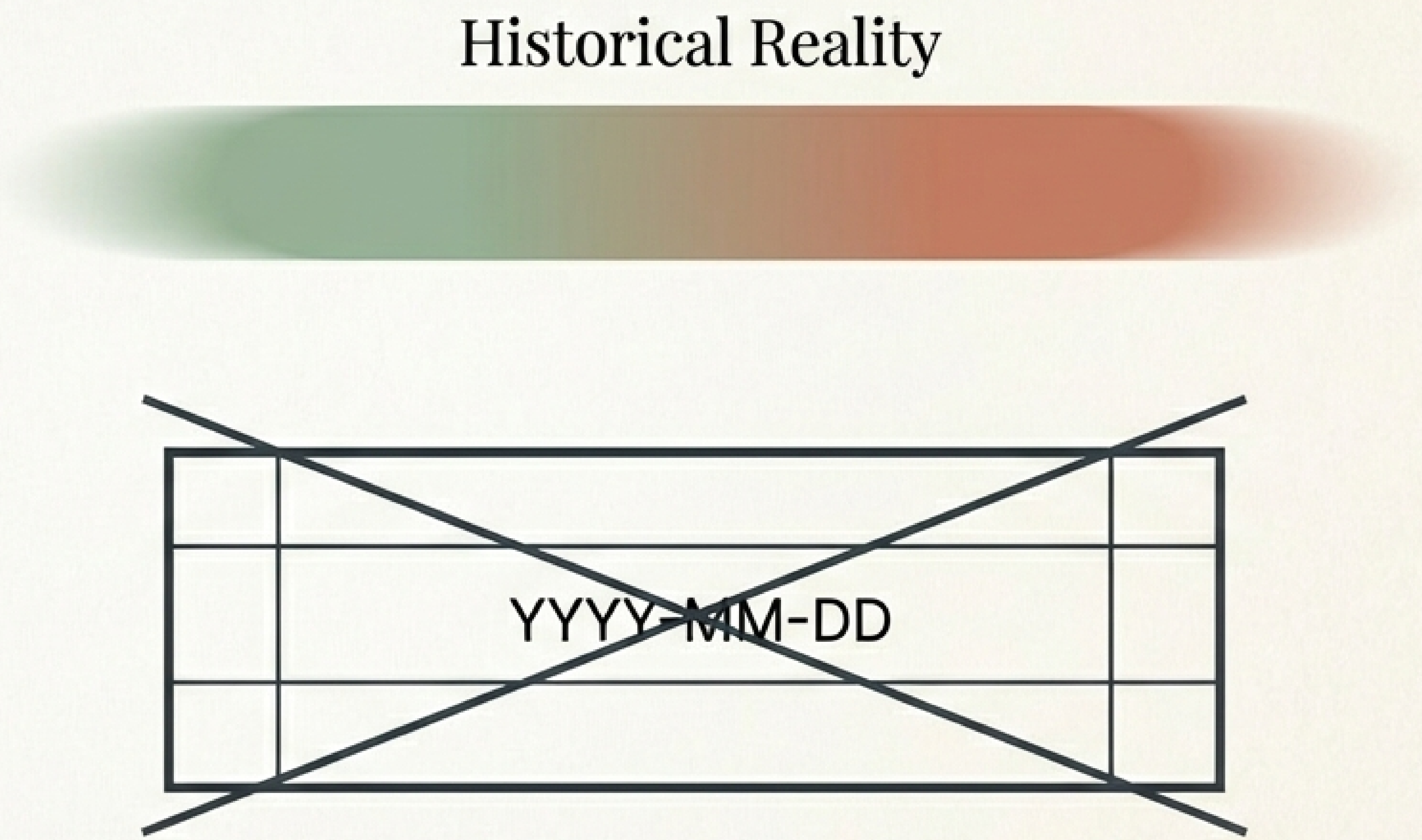
The model captures history as a series of meetings—where persistent items come together and interact in specific Events at specific Places and Times.


E39 Actor
E18 Physical Thing
E5 Event
E53 Place
E52 Time-Span

Embracing Fuzzy Boundaries in Space and Time

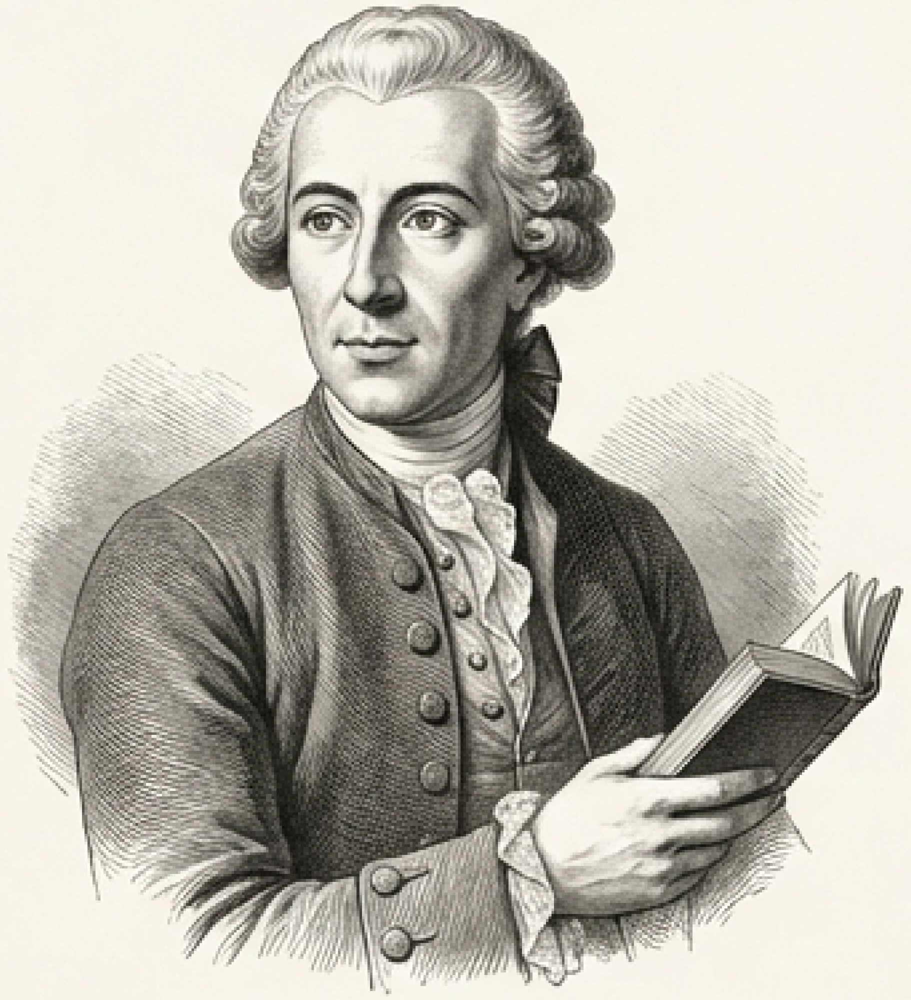
Nature does not provide arbitrarily exact timestamps. Observable historical phenomena have gradual transitions from existing to no longer existing.

The CRM uses E92 Spacetime Volumes and fuzzy temporal boundaries to accurately inter-position events, capturing capturing the inherent nuance and uncerainty of historical data.

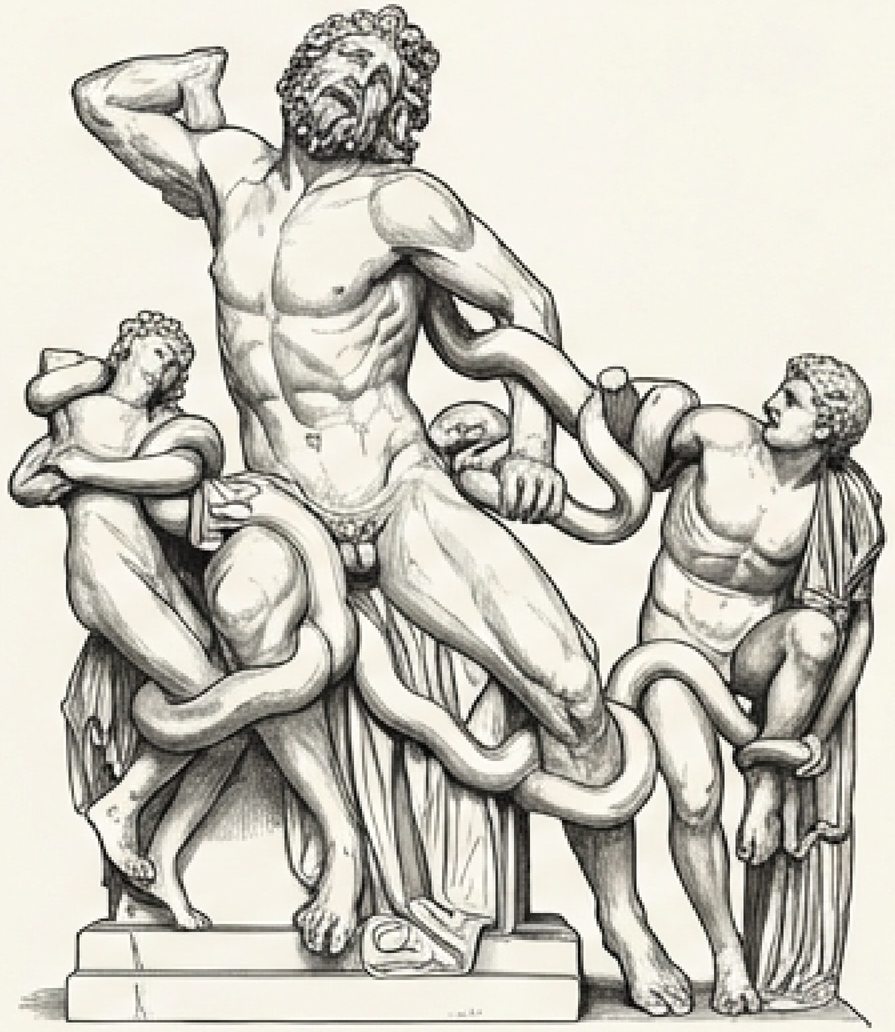


From Theory to Practice: A Historical Encounter

- To see the ontology in action, consider a specific historical fact: In 1755, the German scholar Johann Joachim Winckelmann viewed the Laocoön sculpture at the Vatican in Rome.
- This specific encounter led to his groundbreaking 1764 writing, *History of the Art of Antiquity*.
- How do we formally model this event?



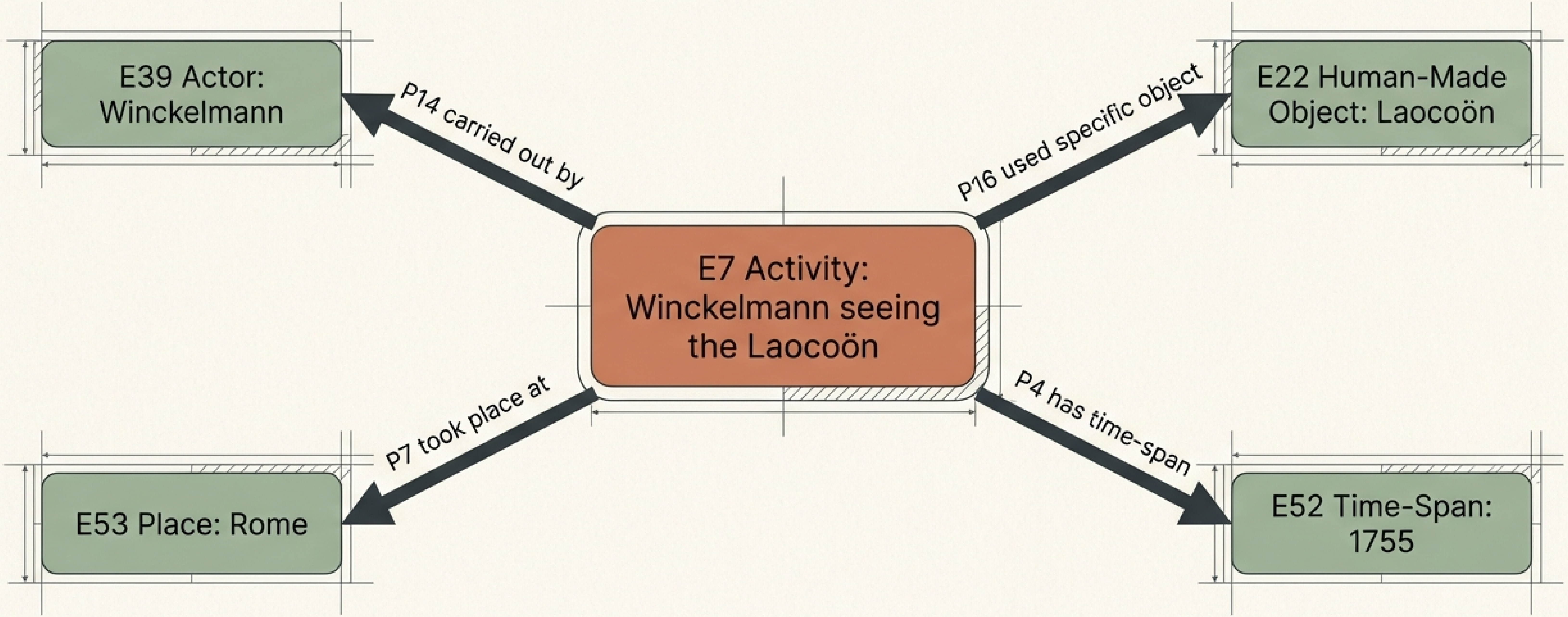
Johann Joachim Winckelmann, 1755



The Laocoön Group, Vatican, Rome

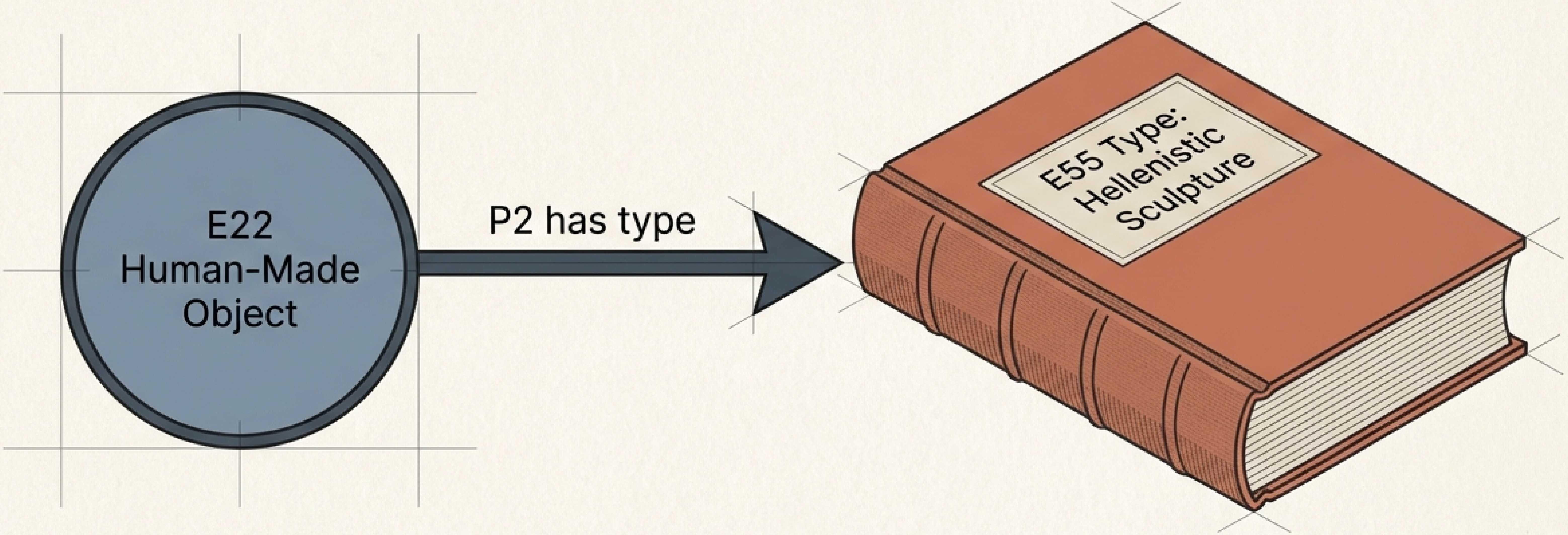
Mapping the Event as a Semantic Graph

By applying formal E-classes and P-properties, an unstructured historical anecdote is transformed into an explicit, machine-readable factual graph.


E39 Actor: Winckelmann
P14 carried out by
P16 used specific object
E22 Human-Made Object: Laocoön
E7 Activity: Winckelmann seeing the Laocoön
P7 took place at
P4 has time-span
E53 Place: Rome
E52 Time-Span: 1755

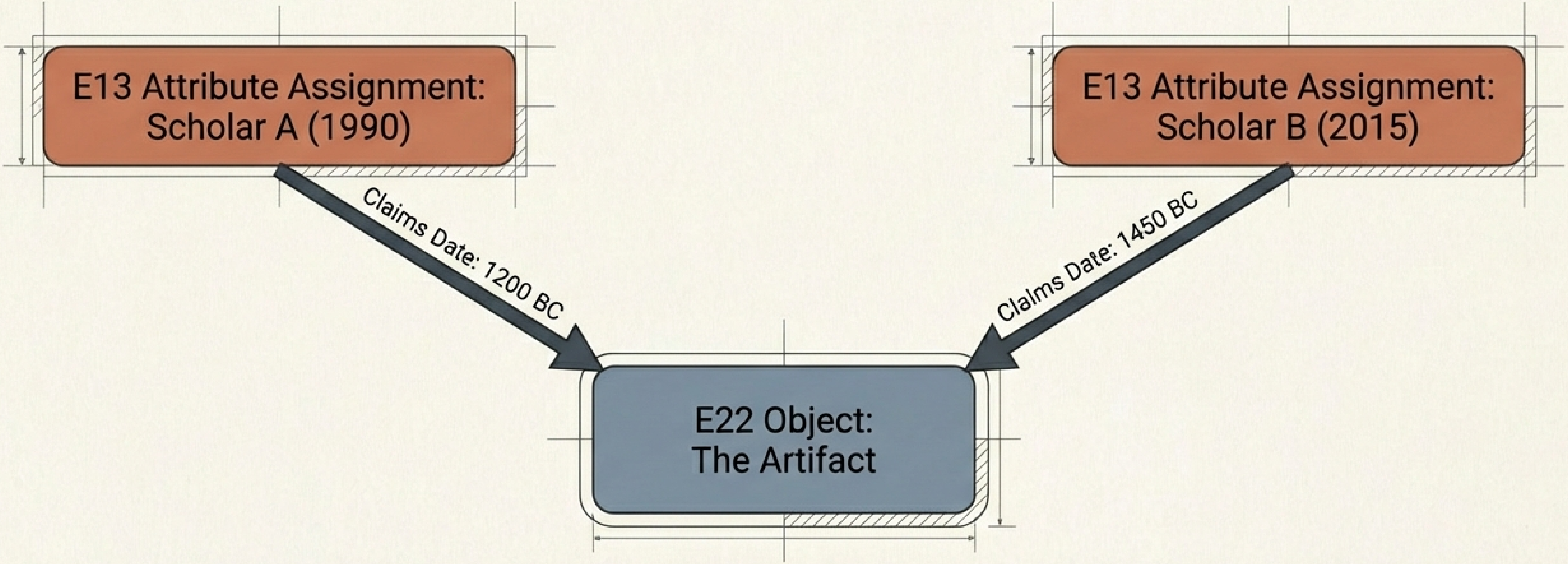
Interfacing with Local Vocabularies

Institutions need to use their own specialized terminologies. The E55 Type class acts as an interface to external thesauri and classification systems. Using the P2 has type property, institutions can apply local categorical knowledge without breaking the underlying structural integrity of the global ontology.


E22 Human-Made Object
P2 has type
E55 Type: Hellenistic Sculpture

Making Space for Conflicting Historical Claims

History is filled with contradictory opinions. The E13 Attribute Assignment class documents exactly who makes a claim and when. This allows conflicting historical attributions to coexist peacefully in the database, attributing authority to the specific agent rather than forcing a single, absolute truth onto the data.



E13 Attribute Assignment:
Scholar A (1990)

E13 Attribute Assignment:
Scholar B (2015)

Claims Date: 1200 BC

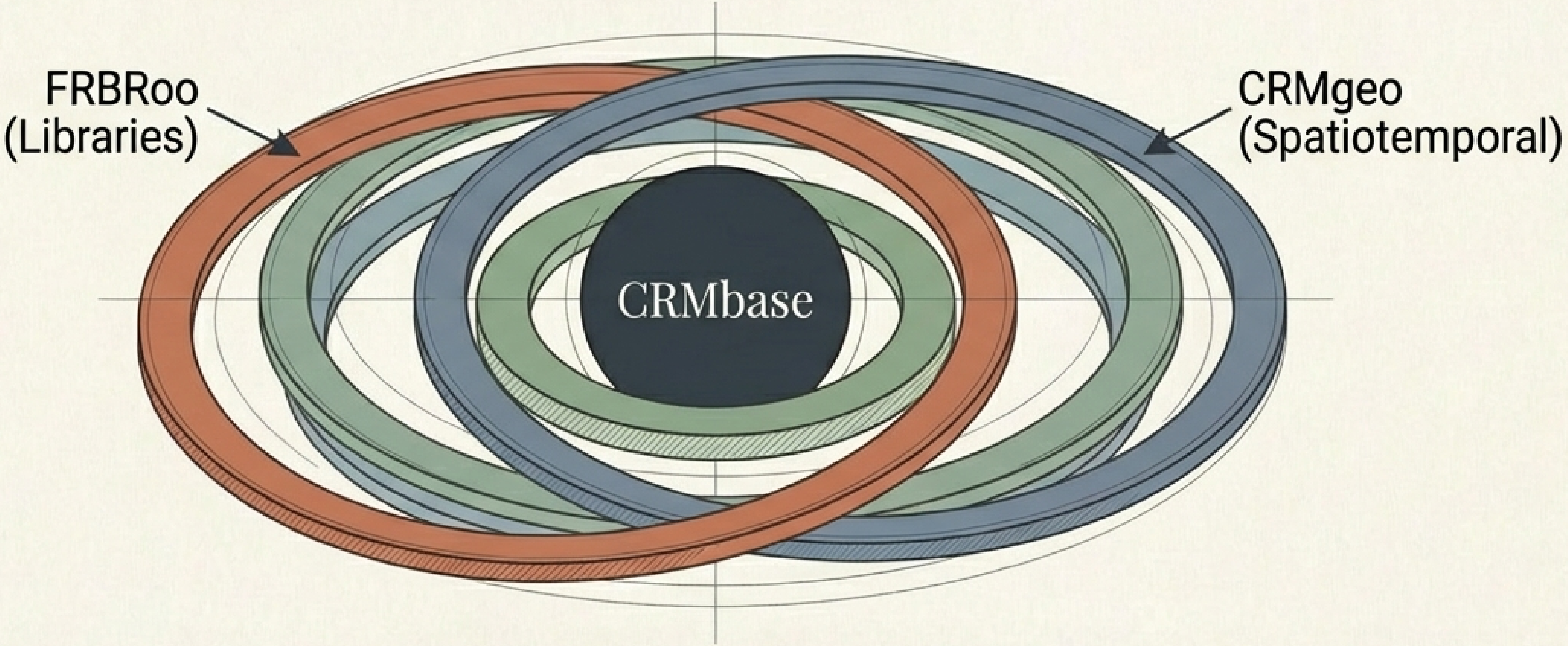
Claims Date: 1450 BC

E22 Object:
The Artifact

Expanding the Scope through Compatible Extensions

The CIDOC CRM is designed to grow. Specific disciplines can build upon the core model through compatible extensions.

Because it operates on the principle of query containment, adding these specialized domains never breaks the base ontology—global queries will still retrieve all relevant facts.


FRBRoo
(Libraries)
CRMgeo
(Spatiotemporal)
CRMbase

A Coherent Global Resource for Human History

The CIDOC CRM is more than a data standard; it is a shared language. By providing a formal, event-centric ontology, it transforms disparate, localized information silos into a coherent, machine-readable, and human-explorable map of the human past.